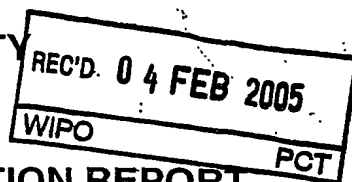


## PATENT COOPERATION TREATY

## PCT

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)





Applicant's or agent's file reference P100526PC00/SJR	<b>FOR FURTHER ACTION</b> See Notification of Transmittal of International Preliminary Examination Report (Form PCT/PEA/416)	
International application No. PCT/GB 03/04854	International filing date (day/month/year) 11.11.2003	Priority date (day/month/year) 11.11.2002
International Patent Classification (IPC) or both national classification and IPC H04L12/56		
Applicant CLEARSPEED TECHNOLOGY PLC et al.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 5 sheets, including this cover sheet.  
  
☒ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).  
  
These annexes consist of a total of 6 sheets.

3. This report contains indications relating to the following items:  

I	<input checked="" type="checkbox"/>	Basis of the opinion
II	<input type="checkbox"/>	Priority
III	<input type="checkbox"/>	Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
IV	<input type="checkbox"/>	Lack of unity of invention
V	<input checked="" type="checkbox"/>	Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
VI	<input type="checkbox"/>	Certain documents cited
VII	<input type="checkbox"/>	Certain defects in the international application
VIII	<input type="checkbox"/>	Certain observations on the international application

Date of submission of the demand  22.03.2004	Date of completion of this report  03.02.2005
Name and mailing address of the international preliminary examining authority:   European Patent Office D-80293 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized Officer  Kreppel, J  Telephone No. +49 89 2399-8246  

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. PCT/GB 03/04854

**I. Basis of the report**

1. With regard to the **elements** of the international application (*Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17)*):

**Description, Pages**

1-3, 6-13 as originally filed  
4, 5, 5a received on 13.12.2004 with letter of 09.12.2004

**Claims, Numbers**

1-25 received on 13.12.2004 with letter of 09.12.2004

**Drawings, Sheets**

1/3-3/3 as originally filed

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language: , which is:

- ☐ the language of a translation furnished for the purposes of the international search (under Rule 23.1(b)).  
☐ the language of publication of the international application (under Rule 48.3(b)).  
☐ the language of a translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☐ contained in the international application in written form.  
☐ filed together with the international application in computer readable form.  
☐ furnished subsequently to this Authority in written form.  
☐ furnished subsequently to this Authority in computer readable form.  
☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.  
☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

- ☐ the description, pages:  
☐ the claims, Nos.:  
☐ the drawings, sheets:

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT**

International application No. **PCT/GB 03/04854**

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)).

*(Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.)*

6. Additional observations, if necessary:

**V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Yes: Claims	1-25
	No: Claims	
Inventive step (IS)	Yes: Claims	1-25
	No: Claims	
Industrial applicability (IA)	Yes: Claims	1-25
	No: Claims	

2. Citations and explanations

**see separate sheet**

**With respect to item V:**

**1 Prior art**

The invention relates to a method of handling packet flows in a communication system (**claim 1**) and a corresponding apparatus (independent **claim 18**). A timestamp is assigned to packets to indicate the order of transmission of the packets. The timestamp is assigned according to scheduling criteria based on a quality of service assigned to a packet flow. The packets have then to be ordered in a queue according to their timestamp. This imposes a high performance requirement when inserting arriving packets into the queue. Document D1=US6396843 addresses this problem by assigning the packet descriptors to a number of bins having different granularities of timestamps. Packets are inserted into these bins according to the required QoS. Packets belonging to flows having e.g. a high rate are assigned to bins with a fine granularity. However, since packets are dequeued from the bins in a LIFO or FIFO order, misordering of packets can occur when dequeuing packets from bins with a coarse granularity. This can cause delay and jitter.

**2 Object**

It is therefore an object of the present invention to provide fast insertion of packets into bins without causing misordering when dequeuing the packets from the bins.

**3 Solution**

This is achieved by a sorting step before outputting a packet. Packets are inserted into coarse bins covering a high range of timestamps when arriving. The bins having a coarse granularity are sorted into a set of bins having a finer granularity. It is to be noted that the sorting step does not completely sort the timestamps but merely splits a bin into ranges. The bins are then repeatedly split up until they are ordered and can be read out to transmit the corresponding packets. The operational complexity of queuing packets according to their timestamp can thus significantly be reduced.

**4 Conclusions**

**INTERNATIONAL PRELIMINARY  
EXAMINATION REPORT - SEPARATE SHEET**

---

International application No. PCT/GB 03/04854

This solution is neither anticipated nor rendered obvious by the prior art at hand.  
The subject-matter of independent **claims 1 and 18** therefore meets the requirements of Article 33 (2) and (3) PCT regarding novelty and inventive step.  
**Claims 2 to 17 and 19 to 25** are dependent on claims 1 and 18 and therefore also novel and inventive.

For example, consider the two popular scheduling methods:

*Fair Queue scheduling* - every packet in the queue is given a finish number, which indicates the relative point in time that the packet is entitled to be outputted. The function that serves packets from the queue must identify the queue whose next packet has the smallest finish number. Ideally, only after the packet has been served and the next packet in the same queue has been revealed can the dequeuing function make its next decision.

*Round Robin scheduling* - Queues are inspected in turn in a predetermined sequence. On each visit a pre-scribed quota of data may be served.

The fundamental problem is how to perform such scheduling algorithms at high speeds. A serialised process can only scale with clock/cycle frequency, or by increasing the depth of the processing pipe which makes the scheduling decision. This approach may only be able to provide a couple of system clock cycles per packet.

On top of this, the scheduling and queue management task is further confounded by a requirement for a large number of potentially very deep queues. Hardware, which executes the scheduling function in a serial manner, is then likely to be highly customised and therefore inflexible if it is to meet the required performance.

US 6,396,843 discloses a packet ordering system in which time-stamped data is queued in a plurality of queues, some of which are subdivided. A two-stage check of the sub-queues first determines which data in each sub-queue has the smallest time-stamp then selects, for transmission, the smallest of those already pre-selected.

### **Summary of the invention**

In its broadest interpretation, the invention, in all its aspects, provides a system for maintaining ordered logical data structures in software at high speeds. The inventive aspects per se are set out below:

In a first aspect, the invention provides a method for handling packet flows, comprising sequences of data packets, in a communication or computer system, the method comprising: assigning an exit number to each said packet; queuing said packets in buffer means; allocating said exit numbers to storage bins, each bin accepting a respective range of orders of exit numbers; and outputting the queued packets in a predetermined order according to an order list determined by said exit numbers assigned to said packets before said packets are queued; the method characterised by the step, before said outputting step, of sorting the contents of a bin containing a first range of exit numbers into a plurality of bins each containing a smaller range of exit numbers.

In a second aspect, the invention provides a data manager for handling packet flows, comprising sequences of data packets, in a communication or computer system, the system comprising: assigning means for assigning an exit number to each said packet; buffer means for queuing said packets; a series of storage bins; allocating means for allocating said exit numbers to said storage bins, each bin accepting a respective range of orders of exit numbers; and output means for outputting the queued packets in a predetermined order according to an order list determined by said exit numbers assigned to said packets before said packets are queued; the system characterised by sorting means upstream of said output means for sorting the contents of a bin containing a first range of exit numbers into a plurality of bins each containing a smaller range of exit numbers.

Exit number information is preferably assigned to packet records which are queued in a separate buffer means to that in which said packets are queued. The packet records are preferably of fixed length and shorter than said packets.

The buffer means for said packet records conveniently consist of said bins, the bins for higher exit number packet records having a larger range than bins for lower exit number packet records.

Where a packet record is assigned an exit number corresponding to a bin that is currently being emptied, that packet record may be held in a specially created secondary bin of the same exit number for emptying after the said bin that is currently being emptied.

The bins may be FIFO buffers, LIFO stacks or a mixture of both.

The queue management is preferably carried out by (a) processing all of the bins in parallel and inserting incoming data into a bin by means of a parallel processor, preferably a SIMD processor, and the bin sorting is preferably carried out by a parallel processor, such as an array processor, preferably a SIMD processor.

#### **Brief Description of the Drawings**

The invention will now be described with reference to the following drawings, in which:

- Figure 1 is a schematic representation of an ideal traffic handling mechanism;
- Figure 2 illustrates the principles of the invention;
- Figure 3 is a functional overview of a system employing the invention;
- Figure 4 shows an implementation of the invention using MTAP processors and state engines.
- Figure 5 is a schematic representation of the way in which order lists are handled; and
- Figure 6 shows how sub-divided order lists are managed.

**Detailed Description of the Illustrated Embodiments*****Introduction***

A third approach to scheduling, in accordance with the invention, is to maintain a single, fully ordered queue instead of multiple FIFO queues. In other words, rather than buffer packets in a set of parallel input queues and then schedule them in some sequence into an output queue, packets are sorted on arrival directly into the output queue.

In comparison with the Fair Queue and Round Robin scheduling approaches, calculations must be made at wire speed for each packet prior to enqueueing but



**Claims:**

1. A method for handling packet flows, comprising sequences of data packets, in a communication or computer system, the method comprising: assigning an exit number to each said packet; queuing said packets in buffer means; allocating said exit numbers to storage bins, each bin accepting a respective range of orders of exit numbers (1, 3-2, 7-4); and outputting the queued packets in a predetermined order according to an order list determined by said exit numbers assigned to said packets before said packets are queued; the method characterised by the step, before said outputting step, of sorting the contents of a bin containing a first range of exit numbers (3-2, 7-4) into a plurality of bins each containing a smaller range of exit numbers (3, 2, 7-6, 5-4).
2. A method as claimed in claim 1, wherein said sorting step is repeated until the contents of the bins are completely sorted.
3. A method as claimed in claim 1 or claim 2, wherein said queuing step comprises placing each said data packet together with its respective exit number in said buffer means, said buffer means comprising said storage bins.
4. A method as claimed in any of claims 1 to 3, wherein said queuing step comprises placing packet records, each record containing information about its respective packet, together with their respective exit numbers in said bins.
5. A method as claimed in claim 4, wherein the packet records are of fixed length.
6. A method as claimed in claim 4 or claim 5, wherein the packet records are shorter than said packets.
7. A method as claimed in any of claims 4 to 6, wherein the bins for higher order exit number packet records have a larger range than bins for lower order exit number packet records.
8. A method as claimed in claim 7, wherein under circumstances in which a packet record is assigned an exit number (2) corresponding to a bin that is currently being emptied, that packet record is held in a specially created secondary bin (2') of the same exit number order for emptying after the said bin that is currently being emptied.

9. A method as claimed in any of claims 4 to 8, wherein the bins are FIFO buffers.
10. A method as claimed in any of claims 4 to 8, wherein the bins are LIFO stacks.
- 5 11. A method as claimed in any of claims 4 to 8, wherein the bins are a mixture of FIFO buffers and LIFO stacks.
- 10 12. A method as claimed in claim 1, wherein queue management is performed by (a) processing all of said bins in parallel and (b) inserting incoming data into a bin by means of a parallel processor.
13. A method as claimed in claim 12, wherein said parallel processor performing said inserting step (b) is an array processor.
- 15 14. A method as claimed in claim 13, wherein said array processor performing said inserting step (b) is a SIMD processor.
15. A method as claimed in claim 1, wherein said sorting step is carried out by a parallel processor.
- 20 16. A method as claimed in claim 15, wherein said parallel processor is an array processor.
- 25 17. A method as claimed in claim 16, wherein said array processor is a SIMD processor.
18. A data manager for handling packet flows, comprising sequences of data packets, in a communication or computer system, the system comprising: assigning means for assigning an exit number to each said packet; buffer means for queuing said packets; a series of storage bins; allocating means for allocating said exit numbers to said storage bins, each bin accepting a respective range of orders of exit numbers (1, 3-2; 7-4); and output means for outputting the queued packets in a predetermined order according to an order list determined by said exit numbers assigned to said packets before said packets are queued; the system characterised by sorting means upstream of said output means for sorting the contents of a bin containing a first range
- 30
- 35

of exit numbers (3-2, 7-4) into a plurality of bins each containing a smaller range of exit numbers (3, 2, 7-6, 5-4).

5 19. A data manager as claimed in claim 18, wherein said allocating means comprises a parallel processor.

20. A data manager as claimed in claim 19, wherein said sorting means comprises a parallel processor.

10 21. A data manager as claimed in claim 19 or claim 20, wherein said parallel processor is an array processor.

22. A data manager as claimed in claim 21, wherein said parallel processor is a SIMD processor.

15

23. A data manager as claimed in any of claims 18 to 22, wherein said buffer means is adapted to queue said data packets together with their respective exit numbers.

20 24. A data manager as claimed in any of claims 18 to 22, wherein said buffer means comprises said series of bins, and wherein said bins are adapted to receive packet records, each record containing information about a respective packet, together with the respective exit number, and said buffer means is adapted to queue said packets.

25 25. A data manager as claimed in claim 18, wherein said sorting means is adapted to repeat sorting said bins until the contents of the bins are completely sorted.